

(19)日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11)特許出願公開番号  
特開2000-101613  
(P2000-101613A)

(43)公開日 平成12年4月7日(2000.4.7)

(51)Int.Cl. <sup>7</sup>	識別記号	F I	テーマコード* (参考)
H 0 4 L 12/40		H 0 4 L 11/00	3 2 1
G 0 6 F 13/00	3 5 3	G 0 6 F 13/00	3 5 3 C
H 0 4 L 12/56		H 0 4 L 11/20	1 0 2 A

審査請求 未請求 請求項の数9 O L (全 14 頁)

(21)出願番号 特願平10-272819

(22)出願日 平成10年9月28日(1998.9.28)

(71)出願人 000005108

株式会社日立製作所

東京都千代田区神田駿河台四丁目6番地

(72)発明者 田口 しほ子

神奈川県川崎市麻生区王禅寺1099番地 株  
式会社日立製作所システム開発研究所内

(72)発明者 岩崎 正明

神奈川県川崎市麻生区王禅寺1099番地 株  
式会社日立製作所システム開発研究所内

(74)代理人 100087170

弁理士 富田 和子

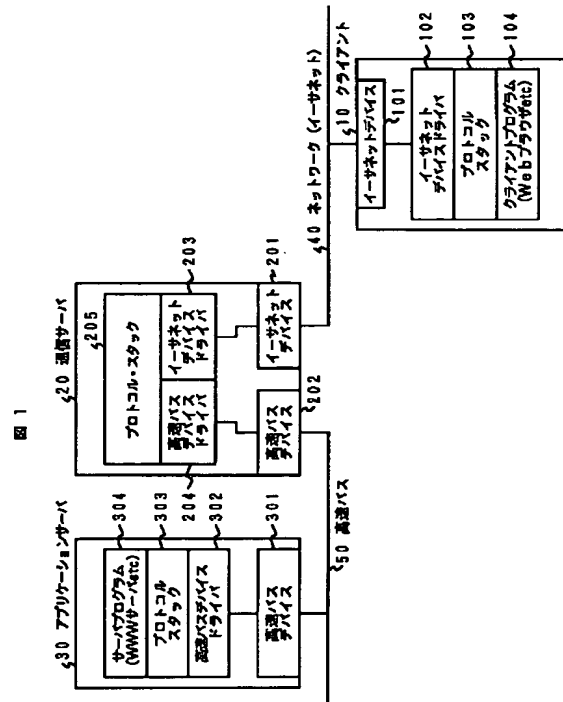
最終頁に続く

(54)【発明の名称】 サーバ・システムおよびクライアント・サーバ間の通信方法

(57)【要約】

【課題】 高スループットネットワークへ対応したクライアント・サーバ・システムを実現する。

【解決手段】 通信サーバ20は、ネットワーク40を介してクライアント10より受信したIPパケット、あるいは当該IPパケットをリアセンブルしたIPパケットをバッファリングし、バッファリングしたデータを高速バス50にて送信可能なサイズでカプセル化して、高速バスパケットとして高速バス50上に送出する。また、高速バス50を介してアプリケーションサーバ30より受信した高速バスパケットからIPパケットを取り出し、ネットワーク40にて送信可能なサイズにフラグメント化して、ネットワーク40上に送出する。



## 【特許請求の範囲】

【請求項 1】クライアントからの要求にしたがい処理を実行するサーバ・システムであって、

ネットワークを介して少なくとも 1 つのクライアントに接続された第 1 のサーバと、前記ネットワークで送信可能な最大パケットサイズよりも大きなサイズのパケットを送信可能なバスを介して前記第 1 のサーバに接続された、前記第 1 のサーバを介して受け取った前記クライアントからの要求にしたがい処理を実行する第 2 のサーバと、を有し、

前記第 1 のサーバは、前記ネットワークを介して前記クライアントより受信したパケットを、当該パケットがフラグメント化されている場合はリアセンブルしてから前記バス上に送出し、フラグメント化されていない場合はそのまま前記バス上に送出するとともに、前記バスを介して前記第 2 のサーバより受信したパケットについて、前記ネットワークにて送信可能なサイズのパケットにフラグメント化して、前記ネットワーク上に送出するパケットサイズ変換手段を備え、

前記第 2 のサーバは、前記バスを介して前記第 1 のサーバより受信したパケットを処理して前記クライアントの要求を復元するとともに、送信すべきデータを前記バスにて送信可能なサイズのパケットにパケット化し、前記バス上に送出するパケット化手段を備えることを特徴とするサーバ・システム。

【請求項 2】請求項 1 記載のサーバ・システムであって、

前記パケットサイズ変換手段は、前記ネットワークを介して前記クライアントより受信した 1 つ以上のパケットを、当該パケットがフラグメント化されている場合はリアセンブルしてから、フラグメント化されていない場合はそのまま、前記バスで送信可能な最大パケットサイズの範囲内で 1 つのパケットにカプセル化し、前記バス上に送出するものであり、

前記パケット化手段は、前記バスを介して受け取った前記カプセル化されたパケットから前記クライアントよりのパケットおよび／または前記クライアントよりのパケットをリアセンブルすることで得られたパケット各々を取り出し、取り出したパケット各々について処理することで前記クライアントの要求を復元することを特徴とするサーバ・システム。

【請求項 3】請求項 2 記載のサーバ・システムであって、

前記パケットサイズ変換手段は、前記受信した 1 つ以上のパケットのデータサイズが所定のサイズに達したときに、前記受信した 1 つ以上のパケットを、当該パケットがフラグメント化されている場合はリアセンブルしてから、フラグメント化されていない場合はそのまま、1 つのパケットにカプセル化して前記バス上に送出することを特徴とするサーバ・システム。

【請求項 4】請求項 2 記載のサーバ・システムであって、

前記パケットサイズ変換手段は、所定時間毎に信号を発生するタイマを備え、前記タイマより信号が出力されたときに、それまでに前記ネットワークを介して前記クライアントより受信した 1 つ以上のパケットを、当該パケットがフラグメント化されている場合はリアセンブルしてから、フラグメント化されていない場合はそのまま、1 つのパケットにカプセル化して前記バス上に送出することを特徴とするサーバ・システム。

【請求項 5】請求項 2 記載のサーバ・システムであって、

前記パケットサイズ変換手段は、所定時間毎に信号を発生するタイマと前記ネットワークを介して受信したパケットの受信間隔を監視する監視手段とを備え、前記監視手段で監視した受信間隔が所定値を越えていない場合は、前記タイマより信号が出力されたときに、それまでに前記ネットワークを介して前記クライアントより受信した 1 つ以上のパケットを、当該パケットがフラグメント化されているときはリアセンブルしてから、フラグメント化されていないときはそのまま、1 つのパケットにカプセル化して前記バス上に送出するとともに、前記監視手段で監視した受信間隔が前記所定値を越えている場合は、前記ネットワークを介して前記クライアントよりパケットを受信する毎に、当該パケットを、フラグメント化されているときはリアセンブルしてから、フラグメント化されていないときはそのまま、1 つのパケットにカプセル化して前記バス上に送出することを特徴とするサーバ・システム。

【請求項 6】請求項 2 記載のサーバ・システムであって、

前記パケットサイズ変換手段は、前記ネットワークを介して前記クライアントより受信した 1 つ以上のパケットを、パケットの送信先あるいは送信元毎に管理し、データサイズが所定のサイズに達したものから順に、当該パケットがフラグメント化されているときはリアセンブルしてから、フラグメント化されていないときはそのまま、1 つのパケットにカプセル化して前記バス上に送出することを特徴とするサーバ・システム。

【請求項 7】請求項 1、2、3、4、5 または 6 記載のサーバ・システムを備えることを特徴とするクライアント・サーバ・システム。

【請求項 8】ネットワークを介して少なくとも 1 つのクライアントに接続されるとともに、前記ネットワークで送信可能な最大パケットサイズよりも大きなサイズのパケットを送信可能なバスを介して、前記クライアントからの要求にしたがい処理を実行するサーバに接続された通信サーバであって、

前記ネットワークを介して前記クライアントより受信したパケットについて、当該パケットがフラグメント化さ

10

20

30

40

50

れている場合はリアセンブルしてから前記バス上に送出し、フラグメント化されていない場合はそのまま前記バス上に送出するとともに、前記バスを介して前記第2のサーバより受信したパケットについて、前記ネットワークにて送信可能なサイズのパケットにフラグメント化して、前記ネットワーク上に送出するパケットサイズ変換手段を備えることを特徴とする通信サーバ。

【請求項9】少なくとも1つのクライアントと、ネットワークを介して前記クライアントに接続された第1のサーバと、前記ネットワークで送信可能な最大パケットサイズよりも大きなサイズのパケットを送信可能なバスを介して前記第1のサーバに接続された、前記第1のサーバを介して受け取った前記クライアントからの要求にしたがい処理を実行する第2のサーバと、を有するクライアント・サーバ・システムにおけるクライアント・サーバ間の通信方法であって、

前記第1のサーバにおいて、前記ネットワークを介して前記クライアントより受信したパケットを、当該パケットがフラグメント化されている場合はリアセンブルしてから前記バス上に送出し、フラグメント化されていない場合はそのまま前記バス上に送出するステップを備え、前記第2のサーバにおいて、前記バスを介して前記第1のサーバより受信したパケットを処理して前記クライアントの要求を取得するステップを備えることを特徴とするクライアント・サーバ間の通信方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、クライアント・サーバ間をネットワークで接続したクライアント・サーバ・システムに関し、特に、クライアント・サーバ間の通信技術に関する。

【0002】

【従来の技術】クライアント・サーバ・システムにおいて、クライアント・サーバ間を繋ぐネットワークとして、イーサネットが広く普及している。イーサネットを用いてデータ転送を行う場合、送信側では、送信データをイーサネットの送信規格である64Byte以上、1.5KByte以下にフラグメント化して送信し、受信側では、フラグメント化されたデータをリアセンブルして、元データを取得する。

【0003】すなわち、送信側の装置において、プロトコル・スタックは、送信を行うアプリケーションからのシステムコールにより、送信データをフラグメント化してパケットを作成するとともに、通信経路のルーティングなどを行い、作成したパケットを送信バッファにコピーしてシステムコールを終了する。

【0004】ここで、イーサネットデバイスは、1つのパケット送信が終了するとパケット送信通知割り込みを発行する。上述したように、プロトコル・スタックは、送信データを64Byte以上、1.5KByte以下

にフラグメント化して送信する。したがって、送信データが1.5KByteより大きかった場合、およそ、送信データの長さを1.5KByteで割った数のパケットを作成していることになる。この場合、プロトコル・スタックは、通常、パケット送信通知割り込みが発生する度にプロセスが中断され、該割り込み処理終了後、残りのパケットを送信する。

【0005】上述した一連の送信処理において、実行中のプロセスから他のプロセスへと処理を移行するために、コンテキスト・スイッチにより、レジスタの待避やロード、各種テーブルやリストなどの更新などが行われる。1回のシステムコールでは、アプリケーションのプロセスからプロトコル・スタックのプロセスへ、およびプロトコル・スタックのプロセスからアプリケーションのプロセスへと、2回のコンテキスト・スイッチが行われる。さらに、パケット送信後の送信通知割り込みにより、実行中のプロセスから割り込み処理へ移行し、その後、元のプロセスへ処理を戻すために、2回のコンテキスト・スイッチが行われる。なお、割り込み処理中にパケット送信処理を行う場合は、割り込み処理中にプロトコル・スタックの処理が加わるため、3回のコンテキスト・スイッチが行われることになる。

【0006】一方、受信側の装置において、イーサネットデバイスは、受信バッファにパケットが到着する度に、パケット到着通知割り込みを発行する。これにより、実行中のプロセスが中断され、割り込み処理が実行される。また、プロトコル・スタックは、この割り込み処理を契機として、到着したパケットをシステムメモリにコピーし、リアセンブルするといった受信処理を行う。受信処理が終了すると、割り込み発生前の処理に戻る。

【0007】上述した一連の受信処理では、実行中のプロセスから割り込み処理へ、割り込み処理からプロトコル・スタックのプロセスへ、およびプロトコル・スタックのプロセスから割り込み処理前に実行していたプロセスへと、3回のコンテキスト・スイッチが行われる。

【0008】

【発明が解決しようとする課題】ところで、近年のネットワーク・スループットの向上により、ネットワーク上の単位時間あたりのデータ転送量が増加している。一方、イーサネットでは、規格により、パケットサイズの最大値が1.5KByte以下に設定されている。このため、サーバにおいて、パケット到着の時間間隔が縮小して単位時間あたりの割り込み発生回数が増加し、これにより、プロセッサの処理能力が不足するといった問題が生じてきている。

【0009】たとえば、一般的な汎用プロセッサ（たとえば、インテル社製のペンティアムプロセッサ）を搭載したコンピュータにおいて、コンテキスト・スイッチの実行には約500～1,500マシンサイクルを要す

10

20

30

40

50

る。この時間は、200MHzの動作周波数を持つプロセッサを搭載したコンピュータでは、約 $2.5\mu\text{sec}$ ～ $7.5\mu\text{sec}$ となる。したがって、上述したように、1回のパケット到着通知割り込みに対して、3回のコンテキスト・スイッチが行われるので、パケットが到着する度に、コンテキスト・スイッチのために約 $7.5\mu\text{sec}$ ～ $22.5\mu\text{sec}$ の時間が費やされることになる。一方、100Mbpsのイーサネットではサイズ64Byte(512bit)のパケットが最大スループットで到着した場合、パケット到着の時間間隔は約 $5\mu\text{sec}$ となる。この場合、パケット到着による割り込み処理時間よりもパケット到着の時間間隔が短くなるため、プロセッサは、パケット到着通知割り込み発行に追従して処理を行うことができなくなる。

【0010】また、送信処理においても、同様に、システムコール発行毎、あるいはパケット送信通知割り込み発行毎に、コンテキスト・スイッチが数回発生するため、プロセッサは、システムコール発行あるいはパケット送信通知割り込み発行に追従して処理を行うことができなくなる。

【0011】さらに、送信処理でのフラグメント化処理やルーティング処理、あるいは、受信処理でのリアセンブル処理などによっても、プロセッサの負荷が増加しており、高スループットネットワークへ対応したクライアント・サーバ・システムの構築が一層困難となっている。

【0012】本発明は、上記事情に基づいてなされたものであり、本発明の目的は、高スループットネットワークへ対応したクライアント・サーバ・システムを実現することにある。特に、既存のネットワークやソフトウェア資産を活用しつつ、サーバにおいて、ネットワーク通信処理にかかるオーバーヘッドを低減することにある。

#### 【0013】

【課題を解決するための手段】上記課題を解決するために、本発明は、クライアントからの要求にしたがい処理を実行するサーバ・システムであって、ネットワークを介して少なくとも1つのクライアントに接続された第1のサーバと、前記ネットワークで送信可能な最大パケットサイズよりも大きなサイズのパケットを送信可能なバスを介して前記第1のサーバに接続された、前記第1のサーバを介して受け取った前記クライアントからの要求にしたがい処理を実行する第2のサーバと、を有し、前記第1のサーバは、前記ネットワークを介して前記クライアントより受信したパケットについて、当該パケットがフラグメント化されている場合はリアセンブルしてから前記バス上に送出し、フラグメント化されていない場合はそのまま前記バス上に送出するとともに、前記バスを介して前記第2のサーバより受信したパケットについて、前記ネットワークにて送信可能なサイズのパケットにフラグメント化して、前記ネットワーク上に送出する

パケットサイズ変換手段を備え、前記第2のサーバは、前記バスを介して前記第1のサーバより受信したパケットを処理して前記クライアントの要求を復元するとともに、送信すべきデータを前記バスにて送信可能なサイズのパケットにパケット化し、前記バス上に送出するパケット化手段を備えることを特徴とする。

【0014】本発明によれば、ネットワークを介してクライアントから送られてきたパケットについて、フラグメント化されているものは、第1のサーバにおいてリアセンブルされた後、第2のサーバに送信される。したがって、第2のサーバが受信するパケットの数を、クライアントが送信したパケット数より減らすことができるので、パケット到着による割り込み発生回数が減少し、クライアントの要求に応じた処理を行う第2のサーバのプロセッサにかかる負荷を低減することができる。

【0015】また、第2のサーバが送信したパケットは、第1のサーバにおいて、ネットワークで使用するサイズのパケット（バスにて送信可能なパケットのサイズよりも小さいサイズのパケット）にフラグメント化された後、ネットワーク上に送出される。したがって、第2のサーバが直接ネットワーク上にパケットを送信する場合に比べて、パケットの送信回数を減らすことができるので、パケット送信による割り込み発生回数が減少し、クライアントの要求に応じた処理を行う第2のサーバのプロセッサにかかる負荷を低減することができる。

【0016】すなわち、本発明によれば、第2のサーバでの通信処理の負担を低減することができ、これにより、高スループットネットワークへ対応したクライアント・サーバ・システムを実現することができる。

【0017】また、本発明によれば、クライアントで使用するアプリケーション（クライアント・プログラム）や第2のサーバで使用するアプリケーション（サーバ・プログラム）、あるいは、ネットワークなどについて、従来のものに対して修正を加える必要がない。このため、既存のソフトウェア資産や既に構築されてあるネットワークを活用しつつ実現することが可能である。

【0018】なお、本発明において、前記パケットサイズ変換手段は、前記ネットワークを介して前記クライアントより受信した1以上のパケットを、当該パケットがフラグメント化されている場合はリアセンブルしてから、フラグメント化されていない場合はそのまま、前記バスで送信可能な最大パケットサイズの範囲内で1つのパケットにカプセル化し、前記バス上に送出するものでもよい。この場合、前記パケット化手段は、前記バスを介して受け取った前記カプセル化されたパケットから前記クライアントよりのパケットおよび／または前記クライアントよりのパケットをリアセンブルすることで得られたパケット各々を取り出し、取り出したパケット各々について処理することで前記クライアントの要求を復元するようにする。

【0019】このようにすることで、ネットワークを介してクライアントから送られてきたパケット、あるいは当該パケットをリアセンブルすることで得られたパケットは、第1のサーバにおいて、複数まとめて1つのパケットとして前記バス上へ送出される。たとえば、ネットワークにてイーサネットが用いられている場合、第1のサーバは、ネットワークを介してクライアントから送られてきたパケットあるいは当該パケットをリアセンブルすることで得られたパケットを、複数まとめて、イーサネットの送信規格で規定されたパケットの最大サイズである1.5 KByteよりも大きなサイズの1つのパケットとして、第2のサーバに送信する。

【0020】したがって、第2のサーバが受信するパケットの数を、クライアントが送信したパケット数より更に減らすことができるので、パケット到着による割り込み発生回数が減少し、クライアントの要求に応じた処理を行う第2のサーバのプロセッサにかかる負荷を更に低減することができる。

#### 【0021】

【発明の実施の形態】以下に、本発明の一実施形態について説明する。

【0022】図1は、本発明の一実施形態が適用されたクライアント・サーバ・システムの概略図である。

【0023】図示するように、クライアント10は、イーサネットなどのネットワーク40を介して、通信サーバ20に接続されている。また、通信サーバ20は、高速バス50を介して、アプリケーションサーバ30と接続されている。ここで、クライアント10、通信サーバ20およびアプリケーションサーバ30は、4つの階層（ftp、HTTPなどのアプリケーション層、TCP、UDPなどのトランスポート層、IPのインターネット層、イーサネット、トークンリングなどのネットワークインターフェース層）に階層化されたプロトコルの集まりであるTCP/IP (Transmission Control Protocol/Internet Protocol) にしたがって通信を行う。ただし、ネットワークインターフェース層については、ネットワーク40で用いるプロトコルで規定されたパケットの最大サイズよりも大きなサイズのパケットを扱うことができ、かつ複数のIPパケットを内部データに含めることが可能なプロトコルが、高速バス50にて用いられている。本実施形態では、ネットワーク40にイーサネットを用いているので、高速バス50では、イーサネットの送信規格で規定されたパケットの最大サイズである1.5 KByteよりも大きなパケットを扱うことができるプロトコル（たとえば、(株)日立製作所のハイパークロスバーネットワーク、Fibre Channel、MyriNet、Tandem社のServerNetなど）が用いられる。

【0024】なお、図1では、1つのクライアント10と1つの通信サーバ20と1つのアプリケーションサーバ

バ30を示しているが、当然のことながら、これらは複数あってもよい。

【0025】クライアント10は、ネットワーク40と接続するためのイーサネットデバイス101を備えたコンピュータにより実現される。すなわち、イーサネットデバイス101を制御するためのイーサネットデバイスドライバ102や、イーサネットデバイス101を介して送受するパケットをTCP/IPにしたがって処理するプロトコル・スタック103や、アプリケーションサーバ30に対して処理を要求するクライアントプログラム（たとえば、Webブラウザプログラムなどのアプリケーションプログラム）104が、ハードディスクドライブなどの記憶装置、あるいは、読取り装置を介してCD-ROM、FDなどの可搬性のある記憶媒体から、メモリ上にロードされ、プロセッサにより実行されることで実現される。

【0026】通信サーバ20は、ネットワーク40と接続するためのイーサネットデバイス201および高速バス50と接続するための高速バスデバイス202を備えたコンピュータにより実現される。すなわち、イーサネットデバイス201を制御するためのイーサネットデバイスドライバ203や、高速バスデバイス202を制御するための高速バスデバイスドライバ204や、ネットワーク40で扱うことができるパケットおよび高速バス50で扱うことができるパケット間の変換処理を行うプロトコル・スタック205が、ハードディスクドライブなどの記憶装置、あるいは、読取り装置を介してCD-ROM、FDなどの可搬性のある記憶媒体から、メモリ上にロードされ、プロセッサにより実行されることで実現される。

【0027】アプリケーションサーバ30は、高速バス50と接続するための高速バスデバイス301を備えたコンピュータにより実現される。すなわち、高速バスデバイス301を制御するための高速バスデバイスドライバ302や、高速バスデバイス301を介して送受するパケットをTCP/IPにしたがって処理するプロトコル・スタック303や、クライアント10の要求を処理するサーバプログラム（たとえば、WWWサーバプログラムなどのアプリケーションプログラム）304が、ハードディスクドライブなどの記憶装置、あるいは、読取り装置を介してCD-ROM、FDなどの可搬性のある記憶媒体から、メモリ上にロードされ、プロセッサにより実行されることで実現される。

【0028】さて、上記構成のクライアント・サーバ・システムにおいて、クライアント10およびアプリケーションサーバ30間の通信は、以下のようにして行われる。

【0029】まず、クライアント10からアプリケーションサーバ30へデータを送信する場合について説明する。

10

20

30

40

50

【0030】図2はクライアント10からアプリケーションサーバ30へデータを送信する場合のパケットの流れを示した図である。

【0031】図示するように、クライアント10において、プロトコル・スタック103は、クライアントプログラム104からのシステムコールにより、クライアントプログラム104が作成した送信データ110をTCP/IPにしたがって処理し、イーサネットの送信規格に適合したサイズのIPパケットを作成する。そして、作成したIPパケットをイーサネットデバイス101の送信バッファ（図示せず）にコピーしてシステムコールを終了する。イーサネットデバイス101は、送信バッファにコピーされたIPパケットを、イーサネットに対応したパケット111としてネットワーク40上に出送する。

【0032】次に、通信サーバ20において、イーサネットデバイス201は、ネットワーク40を介してパケット111を受け取ると、当該パケット111からIPパケットを取り出し、プロトコル・スタック205に渡す。プロトコル・スタック205は、イーサネットデバイス201から受け取ったIPパケットに付されたIPヘッダの情報に基づいて、当該IPパケットがフラグメント化されているか否かを調べる。フラグメント化されている場合はリアセンブルしてからバッファリングし、フラグメント化されていない場合はそのままバッファリングする。上記のバッファリング処理を、バッファリングされたデータサイズが、高速バス50で扱うことができるサイズであって、少なくともイーサネットの送信規格で規定されたパケットの最大サイズである1.5KByteよりも大きなサイズになるまで続け、それから、バッファリングしたデータを、高速バスデバイス202の送信バッファ（図示せず）にコピーする。高速バスデバイス202は、送信バッファにコピーされたデータをカプセル化し、高速バスパケット210として、高速バス50上に出送する。

【0033】次に、アプリケーションサーバ30において、高速バスデバイス301は、高速バス50を介して高速バスパケット210を受け取ると、当該パケット210に含まれる各IPパケット（ネットワーク40を介して通信サーバ20が受け取ったIPパケット、あるいは、当該IPパケットをリアセンブルすることで得られたIPパケット）を取り出し、プロトコル・スタック302に渡す。プロトコル・スタック303は、IPパケットをTCP/IPにしたがって処理し、クライアント10のクライアントプログラム104が作成したデータ110を復元する。サーバプログラム304は、復元したデータ110にしたがいクライアント10の要求に応じた処理を行う。

【0034】ここで、クライアント10からアプリケーションサーバ30へデータを送信する場合における通信

サーバ20の動作をより詳細に説明する。

【0035】図3はクライアント10からアプリケーションサーバ30へデータを送信する場合における通信サーバ20の動作を説明するためのフロー図である。

【0036】まず、イーサネットデバイス201は、ネットワーク40を介してパケットを受信し、イーサネットデバイス201の受信バッファ（図示せず）に格納するとともに、パケット到着通知割り込みを発行する（ステップ1001）。これにより、実行中のプロセスが中断され、割り込み処理が実行される。

【0037】次に、プロトコル・スタック205は、イーサネットデバイス201の受信バッファからIPパケットを通信サーバ20が備えるシステムメモリ（図示せず）にコピーする（ステップ1002）。それから、プロトコル・スタック205は、システムメモリにコピーされているIPパケットのIPヘッダを調べ、当該パケットがフラグメント化されているか否かを判断する（ステップ1003）。そして、フラグメント化されている場合は、フラグメント化された全てのIPパケットが揃っているか否かを調べ（ステップ1003a）、揃っていない場合はステップ1001に戻り、揃っている場合は、フラグメント化されたIPパケットをリアセンブル（ステップ1004）してからシステムメモリにバッファリングする（ステップ1005）。一方、ステップ1003において、フラグメント化されていない場合は、当該IPパケットをそのままシステムメモリにバッファリングする（ステップ1005）。

【0038】次に、プロトコル・スタック205は、システムメモリにバッファリングされているデータが、高速バス50で扱うことができるパケットのサイズであって、少なくともイーサネットの送信規格で規定されたパケットの最大サイズである1.5KByteよりも大きなサイズに達しているか否かを判断し（ステップ1006）、達している場合はステップ1007に移行し、達していない場合は、ステップ1001に戻って、イーサネットデバイス201に次のIPパケットが到着するのを待つ。

【0039】次に、プロトコル・スタック205は、システムメモリにバッファリングされているデータが、前記パケットのサイズに達している場合、当該データを高速バスデバイス202の送信バッファにコピーする（ステップ1007）。これを受けて、高速バスデバイス202は、送信バッファにコピーされたデータをカプセル化し、高速バスパケットとして、高速バス50上に出送する（ステップ1008）。

【0040】これにより、クライアント10が送信したデータは、イーサネットの規格に適合したサイズのパケットでネットワーク40上を伝送され、通信サーバ20において、イーサネットの規格サイズよりも大きなサイズに変換されて、高速バス50を経てアプリケーション

サーバ 30 に伝送される。

【0041】次に、アプリケーションサーバ 30 からクライアント 10 ヘデータを送信する場合について説明する。

【0042】図 4 はアプリケーションサーバ 30 からクライアント 10 ヘデータを送信する場合のパケットの流れを示した図である。

【0043】図示するように、アプリケーションサーバ 30 において、プロトコル・スタック 303 は、サーバプログラム 304 からのシステムコールにより、サーバプログラム 304 が作成した送信データ 310 を TCP / IP にしたがって処理し、高速バス 50 で扱うことができるサイズの IP パケットを作成する。そして、作成した IP パケットを高速バスデバイス 301 の送信バッファ（図示せず）にコピーしてシステムコールを終了する。高速バスデバイス 301 は、送信バッファにコピーされた IP パケットをカプセル化し、高速バスパケット 311 として、高速バス 50 上に送出する。

【0044】次に、通信サーバ 20 において、高速バスデバイス 202 は、高速バス 50 を介して高速バスパケット 311 を受け取ると、当該パケット 311 から IP パケットを取り出し、プロトコル・スタック 205 に渡す。プロトコル・スタック 205 は、受け取った IP パケットを、ネットワーク 40 で扱うことができるサイズ（ネットワーク 40 で用いるイーサネットの送信規格で規定されたパケットの最大サイズである 1.5 KByte よりも小さなサイズ）にフラグメント化し、IP パケットを作成する。そして、作成した IP パケットをイーサネットデバイス 201 の送信バッファ（図示せず）にコピーする。イーサネットデバイス 201 は、送信バッファにコピーされたパケットを、イーサネットに対応したパケット 211 にして、順次、ネットワーク 40 上に送出する。

【0045】次に、クライアント 10 において、イーサネットデバイス 101 は、ネットワーク 40 を介して受け取ったパケット 211 から IP パケットを取り出し、プロトコル・スタック 103 に渡す。プロトコル・スタック 103 は、IP パケットを TCP / IP にしたがって処理し、アプリケーションサーバ 30 のサーバプログラム 304 が作成したデータ 310 を復元する。クライアントプログラム 104 は、復元したデータ 310 にしたがって処理を行う。

【0046】ここで、アプリケーションサーバ 30 からクライアント 10 ヘデータを送信する場合における通信サーバ 20 の動作をより詳細に説明する。

【0047】図 5 はアプリケーションサーバ 30 からクライアント 10 ヘデータを送信する場合における通信サーバ 20 の動作を説明するためのフロー図である。

【0048】まず、高速バスデバイス 301 は、高速バス 50 を介して高速バスパケットを受信し、高速バスデ

バイス 301 の受信バッファ（図示せず）に格納するとともに、パケット到着通知割り込みを発行する（ステップ 2001）。これにより、実行中のプロセスが中断され、割り込み処理が実行される。

【0049】次に、プロトコル・スタック 205 は、高速バスデバイス 301 の受信バッファから IP パケットを通信サーバ 20 が備えるシステムメモリ（図示せず）にコピーする（ステップ 2002）。次に、プロトコル・スタック 205 は、システムメモリにコピーされている IP パケットを、ネットワーク 40 で扱うことができるパケットサイズ（イーサネットの送信規格で規定されたパケットの最大サイズである 1.5 KByte よりも小さなサイズ）にフラグメント化し、IP パケットを作成する（ステップ 2003）。次に、プロトコル・スタック 205 は、作成した IP パケットをイーサネットデバイス 201 の送信バッファにコピーする。イーサネットデバイス 201 は、送信バッファにコピーされた IP パケットを、イーサネットに対応したパケットにしてネットワーク 40 上に送出する（ステップ 2004）。ここで、イーサネットデバイス 40 は、パケットを 1 つ送出する毎に送信通知割り込み通知を発行する（ステップ 2005）。

【0050】プロトコル・スタック 205 は、イーサネットデバイス 201 から送信通知割り込み通知を受け取ると、送信バッファに未送信の IP パケットが残っているか否かを調べ（ステップ 2006）、残っている場合は、ステップ 2004 に戻って残りのパケットをネットワーク 40 上に送出する。

【0051】これにより、アプリケーションサーバ 30 が送信したデータは、高速バス 50 の規格に適合したサイズのパケットにフラグメント化されて高速バス 50 上に伝送され、通信サーバ 20 において、イーサネットの規格サイズよりも小さなサイズに変換されて、ネットワーク 40 を経てクライアント 10 に伝送される。

【0052】以上、本発明の一実施形態について説明した。

【0053】本実施形態によれば、ネットワーク 40 を介してクライアント 10 から送られてきたパケットは、通信サーバ 20 において、前記パケットよりも大きなサイズのパケットに変換されて、アプリケーションサーバ 30 に送信される。したがって、アプリケーションサーバ 30 が受信するパケットの数を、クライアント 10 が送信したパケット数より減らすことができるので、パケット到着による割り込み発生回数が減少し、アプリケーションサーバ 30 のプロセッサにかかる負荷を低減することができる。

【0054】また、アプリケーションサーバ 30 が送信したパケットは、通信サーバ 20 において、ネットワーク 40 で使用する、前記パケットよりも小さなサイズのパケットに変換されて、ネットワーク 40 上に送出され

10

20

30

40

50

る。したがって、アプリケーションサーバ 30 が直接ネットワーク 40 上にパケットを送信する場合に比べて、パケットの送信回数を減らすことができるので、パケット送信による割り込み発生回数が減少し、アプリケーションサーバ 30 のプロセッサにかかる負荷を低減することができる。

【0055】すなわち、アプリケーションサーバ 30 での通信処理の負担を低減することができ、これにより、高スループットネットワークへ対応したクライアント・サーバ・システムを実現することができる。

【0056】また、本発明によれば、クライアント 10 で使用するクライアントプログラム 104 やアプリケーションサーバ 30 で使用するサーバプログラム 304、クライアント 10 およびアプリケーションサーバ 30 で使用するプロトコル・スタック 103、303、あるいは、ネットワーク 40 などについて、従来のものに対して修正を加える必要がない。このため、既存のソフトウェア資産や既に構築されてあるネットワークを活用しつつ実現することが可能である。

【0057】なお、本発明は上記の実施形態に限定されるものではなく、その要旨の範囲内で数々の変形が可能である。

【0058】たとえば、上記の実施形態では、通信サーバ 20 において、ネットワーク 40 を介して受け取った IP パケット、あるいは当該 IP パケットをリアセンブルした IP パケットをシステムメモリにバッファリングし、バッファリングしたデータが高速バス 50 で扱うことができるサイズ（イーサネットの送信規格で規定されたパケットの最大サイズである 1.5 KByte よりも大きなサイズ）に達したときに、当該バッファリングしたデータをカプセル化し、高速バスパケットとして、高速バス 40 上に送出するようにしている。

【0059】しかしながら、本発明はこれに限定されるものではない。たとえば、通信サーバ 20 に、所定時間毎に信号を発生するタイマを設け、IP パケットをタイマより信号が出力されるまでシステムメモリにバッファリングし、タイマより信号が出力されたときに、バッファリングしたデータを、高速バス 50 で扱うことができるパケットサイズでカプセル化し、高速バスパケットとして、高速バス 40 上に送出するようにしてもよい。なお、タイマが信号を発生する時間間隔は、使用するネットワーク 40 のスループットや、アプリケーションサーバや、クレーム 10 上のアプリケーションの要求スループットなどに基づいて求めることができる。たとえば、システムメモリにバッファリングしたデータが高速バス 50 で扱うことができるパケットサイズに達するのに要する時間よりも長くなるように設定する。この時間は、使用するネットワーク 40 のスループットに基づいて求めることができる。

【0060】図 6 は、上記のように修正した本発明の一

実施形態における、クライアント 10 からアプリケーションサーバ 30 へデータを送信する場合における通信サーバ 20 の動作を説明するためのフロー図である。

【0061】まず、イーサネットデバイス 201 は、ネットワーク 40 を介してパケットを受信し、イーサネットデバイス 201 の受信バッファに格納するとともに、パケット到着通知割り込みを発行する（ステップ 3001）。これにより、実行中のプロセスが中断され、割り込み処理が実行される。

10 【0062】次に、プロトコル・スタック 205 は、イーサネットデバイス 201 の受信バッファから IP パケットを通信サーバ 20 が備えるシステムメモリ（図示せず）にコピーする（ステップ 3002）。それから、プロトコル・スタック 205 は、システムメモリにコピーされている IP パケットの IP ヘッダを調べ、当該パケットがフラグメント化されているか否かを判断する（ステップ 3003）。そして、フラグメント化されている場合は、フラグメント化された全ての IP パケットが揃っているか否かを調べ（ステップ 3003a）、揃っていない場合はステップ 3001 に戻り、揃っている場合は、フラグメント化された IP パケットをリアセンブル（ステップ 3004）してからシステムメモリにバッファリングする（ステップ 3005）。一方、ステップ 3003 において、フラグメント化されていない場合は、当該 IP パケットをそのままシステムメモリにバッファリングする（ステップ 3005）。

20 【0063】次に、プロトコル・スタック 205 は、通信サーバに設けられたタイマ（図示せず）より信号が出力されたか否かを判断し（ステップ 3006）、出力された場合はステップ 3007 に移行し、出力されていない場合はステップ 3001 に戻って、イーサネットデバイス 201 に次のパケットが到着するのを待つ。

30 【0064】次に、プロトコル・スタック 205 は、タイマより信号が出力された場合、バッファリングされているデータのうち、高速バス 50 で扱うことができるサイズであって、少なくとも、イーサネットの送信規格で規定されたパケットの最大サイズである 1.5 KByte よりも大きなサイズ分のデータを高速バスデバイス 202 の送信バッファにコピーする（ステップ 3007）。これを受けて、高速バスデバイス 202 は、送信バッファにコピーされたデータをカプセル化し、高速バスパケットとして、高速バス 50 上に送出する（ステップ 3008）。ここで、高速バスデバイス 202 は、高速バスパケットを 1 つ送出する毎に送信通知割り込み通知を発行する（ステップ 3009）。

40 【0065】プロトコル・スタック 205 は、高速バスデバイス 202 から送信通知割り込み通知を受け取ると、システムメモリに未送信のデータが残っているか否かを調べ（ステップ 3010）、残っている場合は、ステップ 3007 に戻って残りのデータを高速バスパケッ



トとしてネットワーク４０上に出送する。

【００６６】これにより、クライアント１０が送信したデータは、イーサネットの規格に適合したサイズにパケット化されてネットワーク４０上を伝送され、通信サーバ２０において、高速バス５０の規格に適合した、イーサネットの規格サイズよりも大きなサイズのパケットに変換されて、高速バス５０を経てアプリケーションサーバ３０に伝送される。

【００６７】さらに、本実施形態において、通信サーバ２０に、所定時間毎に信号を発生するタイマと、ネットワーク４０上を伝送するパケットのイーサネットデバイス２０１への到着時間間隔を計測するパケット到着間隔計測部を設け、パケット到着間隔計測部で監視した時間間隔が所定値以下の場合、上記説明した図６に示すフローにしたがい、タイマより信号が出力されたときにシステムメモリにバッファリングされたデータをパケット化して高速バス４０上に出送し、所定値を越えている場合は、図７に示すフローにしたがい、イーサネットデバイス２０１からパケット到着通知割り込みが発行される毎に、システムメモリにバッファリングされたデータをパケット化して高速バス４０上に出送するようにしてもよい。

【００６８】なお、パケット到着間隔計測部は、イーサネットデバイス２０１から通知されるパケット到着割り込みの通知間隔を計測することで実現できる。また、前記所定値は、システムメモリに割り当てられた、データをバッファリングするバッファのサイズ、ネットワーク４０上を伝送するパケットのイーサネットデバイス２０１への到着時間間隔とバッファリングされるデータのサイズの変化との関係などを考慮して、設定すればよい。

【００６９】ここで、図７に示すフローについて説明する。

【００７０】まず、イーサネットデバイス２０１は、ネットワーク４０を介してパケットを受信し、イーサネットデバイス２０１の受信バッファに格納するとともに、パケット到着通知割り込みを発行する（ステップ５００１）。これにより、実行中のプロセスが中断され、割り込み処理が実行される。

【００７１】次に、プロトコル・スタック２０５は、イーサネットデバイス２０１の受信バッファからＩＰパケットを通信サーバ２０が備えるシステムメモリにコピーする（ステップ５００２）。それから、プロトコル・スタック２０５は、システムメモリにコピーされているＩＰパケットのＩＰヘッダを調べ、当該パケットがフラグメント化されているか否かを判断する（ステップ５００３）。そして、フラグメント化されている場合は、フラグメント化された全てのＩＰパケットが揃っているか否かを調べ（ステップ５００３ａ）、揃っていない場合はステップ１００１に戻り、揃っている場合は、フラグメント化されたＩＰパケットをリアセンブル（ステップ５

００４）してから高速バスデバイス２０２の送信バッファにコピーする（ステップ５００５）。一方、ステップ５００３において、フラグメント化されていない場合は、当該ＩＰパケットをそのまま高速バスデバイス２０２の送信バッファにコピーする（ステップ５００５）。これを受けて、高速バスデバイス２０２は、送信バッファにコピーされたデータをカプセル化し、高速バスパケットとして、高速バス５０上に出送する（ステップ５００６）。

【００７２】これにより、クライアント１０が送信したデータは、イーサネットの規格に適合したサイズのパケットでネットワーク４０上を伝送され、通信サーバ２０においてはリアセンブルされてから、高速バス５０を経てアプリケーションサーバ３０に伝送される。

【００７３】また、上記の実施形態では、ネットワーク４０に接続されている通信サーバ２０が１つの場合について説明したが、当然のことながら、複数の通信サーバ２０がネットワーク４０に接続されるように構成してもよい。

【００７４】さらに、上記の実施形態では、高速バス５０に接続されているアプリケーションサーバ３０が１つの場合について説明したが、当然のことながら、複数のアプリケーションサーバ３０が高速バス５０に接続されるように構成してもよい。また、各アプリケーションサーバ３０で稼働するサーバプログラム３０４は複数のポートを使用するものであってもよい。

【００７５】この場合、通信サーバ２０において、イーサネットデバイス２０１を介して受け取ったＩＰパケット、あるいは当該ＩＰパケットをリアセンブルすることで得られたＩＰパケットをシステムメモリにバッファリングするに際し、前記ＩＰパケットをＩＰヘッダの情報により特定される送信先（アプリケーションサーバ３０単位、あるいはサーバプログラム３０４のポート単位）毎にバッファリングするようにしてもよい。そして、送信先毎にバッファリングされたデータのサイズをモニタし、当該サイズが高速バス５０で扱えるパケットサイズに達したものをカプセル化し、高速バスパケットとして、高速バス５０に出送するようにしてもよい。

【００７６】図８は、上記のように修正した本発明の一実施形態における、クライアント１０からアプリケーションサーバ３０へデータを送信する場合における通信サーバ２０の動作を説明するためのフロー図である。

【００７７】まず、イーサネットデバイス２０１は、ネットワーク４０を介してパケットを受信し、イーサネットデバイスの受信バッファに格納するとともに、パケット到着通知割り込みを発行する（ステップ４００１）。これにより、実行中のプロセスが中断され、割り込み処理が実行される。

【００７８】次に、プロトコル・スタック２０５は、イ

イーサネットデバイス 201 の受信バッファから IP パケットを通信サーバ 20 が備えるシステムメモリ（図示せず）にコピーする（ステップ 4002）。それから、プロトコル・スタック 205 は、システムメモリにコピーされている IP パケットの IP ヘッダを調べ、当該パケットがフラグメント化されているか否かを判断する（ステップ 4003）。そして、フラグメント化されている場合は、フラグメント化された全ての IP パケットが揃っているか否かを調べ（ステップ 4003a）、揃っていない場合はステップ 4001 に戻り、揃っている場合は、フラグメント化された IP パケットをリアセンブル（ステップ 4004）してからステップ 4005 に移行する。一方、ステップ 4003 において、フラグメント化されていない場合は、直ちにステップ 4005 に移行する。

【0079】ステップ 4005 において、プロトコル・スタック 205 は、ネットワーク 40 を介して受け取った IP パケット、あるいは、当該 IP パケットをリアセンブルすることで得られた IP パケットの IP ヘッダを調べ、送信先のサーバプログラム 304 のポートを取得する（ステップ 4005）。そして、IP パケットを、システムメモリに、前記取得した送信先のサーバプログラム 304 のポート毎にバッファリングする（ステップ 4006）。

【0080】次に、プロトコル・スタック 205 は、サーバプログラム 304 のポート毎にバッファリングされたデータのうち、データサイズが高速バス 50 で扱うことができるパケットのサイズに達しているものがあるか否かを判断する（ステップ 4007）。達しているものがある場合はステップ 4008 に移行し、達しているものがない場合はステップ 4001 に戻って、イーサネットデバイス 201 に次のパケットが到着するのを待つ。

【0081】プロトコル・スタック 205 は、サーバプログラム 304 のポート毎にバッファリングされたデータのうち、データサイズが高速バス 50 で扱うことができるパケットサイズであって、少なくともイーサネットの送信規格で規定されたパケットの最大サイズである

1. 5 KByte よりも大きなサイズに達しているものがある場合、当該データを高速バスデバイス 202 の送信バッファにコピーする（ステップ 4008）。これを受けて、高速バスデバイス 202 は、送信バッファにコピーされたデータをカプセル化し、高速バスパケットとして、高速バス 50 上に送出する（ステップ 4009）。ここで、高速バスデバイス 202 は、高速バスパケットを送出すると送信通知割り込み通知を発行する。

【0082】プロトコル・スタック 205 は、高速バスデバイス 202 から送信通知割り込み通知を受け取ると、システムメモリにバッファリング・データが残っているか否かを調べ（ステップ 4008）、残っている場合は、ステップ 4001 に戻って、イーサネットデバイ

ス 201 に次のパケットが到着するのを待つ。

【0083】このようにすることで、ネットワーク 40 を介してクライアント 10 から受け取った IP パケットは、IP ヘッダの情報により特定される送信先単位でまとめられて、アプリケーションサーバ 30 へ送信される。

【0084】なお、図 8 に示すフローでは、イーサネットデバイス 201 を介して受け取った IP パケット、あるいは、当該 IP パケットをリアセンブルすることで得られた IP パケットをシステムメモリにバッファリングするに際し、IP ヘッダの情報により特定される送信先毎にバッファリングするものについて説明しているが、IP ヘッダの情報により特定される送信元（クライアント 10 単位、あるいはクライアントプログラム 104 のポート単位）毎にバッファリングするように修正することも可能である。

【0085】また、上記の実施形態では、イーサネットデバイス 201 を介して受け取った IP パケットを一旦システムメモリにコピーし、それから、IP パケットを、フラグメント化されているものはリアセンブルしてから、フラグメント化されていないものはそのまま、システムメモリにバッファリングし、バッファリングされた IP パケットのデータサイズが所定のサイズに達したときに、あるいはタイマから信号が出力されたときに、前記バッファリングされた IP パケットを 1 つのパケットにカプセル化して高速バスパケットとし、高速バス 50 上に送出するものについて説明した。しかしながら、本発明はこれに限定されるものではない。

【0086】たとえば、イーサネットデバイス 201 を介して受け取った IP パケットをシステムメモリにコピーし、コピーした IP パケットのデータサイズが所定のサイズに達したときに、あるいはタイマから信号が出力されたときに、前記コピーした IP パケットを、フラグメント化されているものはリアセンブルしてから、フラグメント化されていないものはそのまま、1 つのパケットにカプセル化して高速バスパケットとし、高速バス 50 上に送出するようにしてもよい。

【0087】あるいは、イーサネットデバイス 201 からパケット到着通知割り込みが発行される毎に、イーサネットデバイス 201 を介して受け取った IP パケットを、フラグメント化されているものはリアセンブルしてから、フラグメント化されていないものはそのまま、カプセル化して高速バスパケットとし、高速バス 50 上に送出するようにしてもよい。この場合でも、ネットワーク 40 上でフラグメント化されている IP パケットがリアセンブルされて高速バス 50 上に送出されることになるので、アプリケーションサーバ 30 が受信するパケット数を減らすことができる。

【0088】また、本実施形態では、ネットワークインターフェース層として、ネットワーク 40 にイーサネッ

10

20

30

40

50

19

トを用いた場合について説明したが、その他のプロトコル、たとえばトークリングを用いた場合でも、高速バスで用いるパケットのサイズをトークリングで規定するパケットサイズよりも大きくすることで、本発明は同様に実現できる。

#### 【0089】

【発明の効果】以上説明したように、本発明によれば、通信処理にかかるサーバの負担を分散させることができ、これにより、高スループットネットワークへ対応したクライアント・サーバ・システムを実現することができる。

#### 【図面の簡単な説明】

【図1】本発明の一実施形態が適用されたクライアント・サーバ・システムの概略図である。

【図2】クライアント10からアプリケーションサーバ30へデータを送信する場合のパケットの流れを示した図である。

【図3】クライアント10からアプリケーションサーバ30へデータを送信する場合における通信サーバ20の動作を説明するためのフロー図である。

【図4】アプリケーションサーバ30からクライアント10へデータを送信する場合のパケットの流れを示した図である。

【図5】アプリケーションサーバ30からクライアント10へデータを送信する場合における通信サーバ20の動作を説明するためのフロー図である。

20

【図6】本発明の一実施形態の変形例における、クライアント10からアプリケーションサーバ30へデータを送信する場合における通信サーバ20の動作を説明するためのフロー図である。

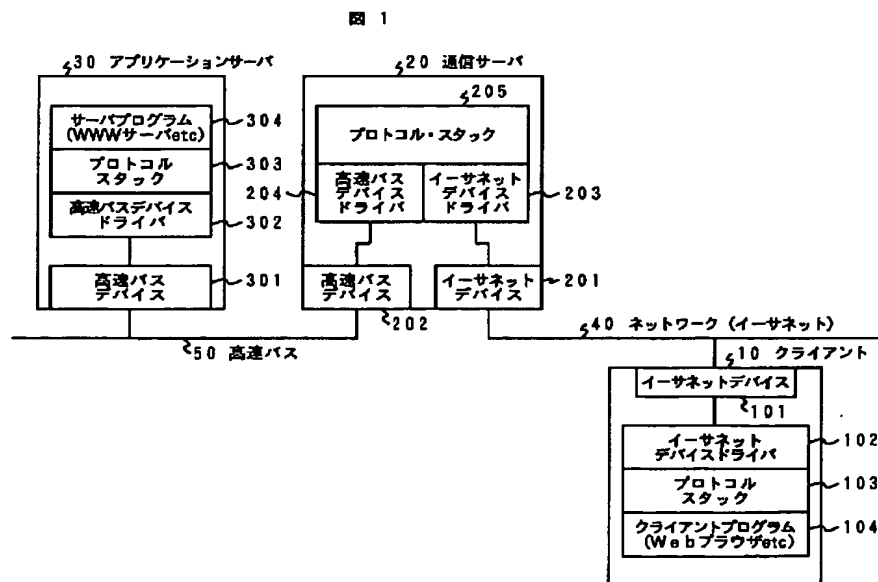
【図7】本発明の一実施形態の変形例における、クライアント10からアプリケーションサーバ30へデータを送信する場合における通信サーバ20の動作を説明するためのフロー図である。

【図8】本発明の一実施形態の変形例における、クライアント10からアプリケーションサーバ30へデータを送信する場合における通信サーバ20の動作を説明するためのフロー図である。

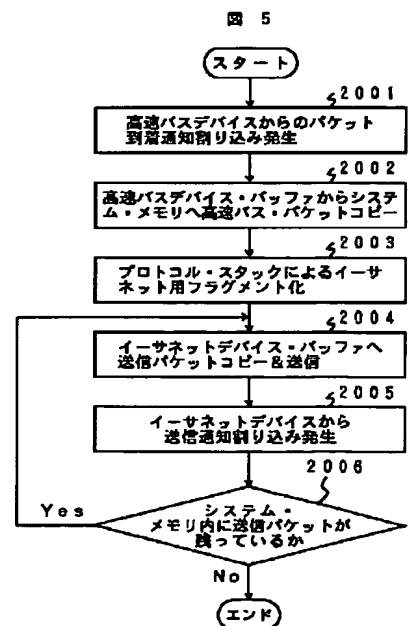
#### 【符号の説明】

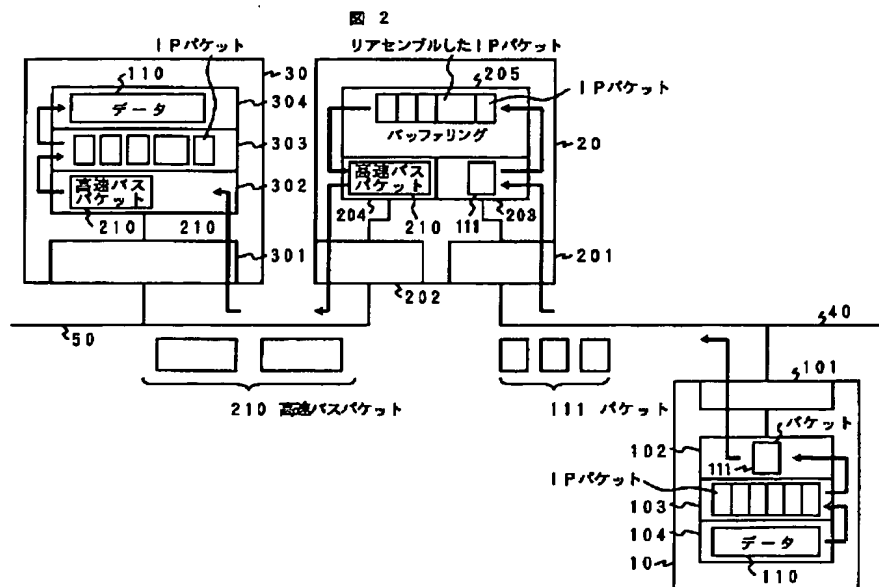
- 10 クライアント
- 20 通信サーバ
- 30 アプリケーションサーバ
- 40 ネットワーク
- 50 高速バス
- 101、201 イーサネットデバイス
- 102、203 イーサネットデバイスドライバ
- 103、205、303 プロトコル・スタック
- 104 クライアントプログラム
- 111、211 パケット
- 202、301 高速バスデバイス
- 210、311 高速バスパケット
- 304 サーバプログラム

【図1】

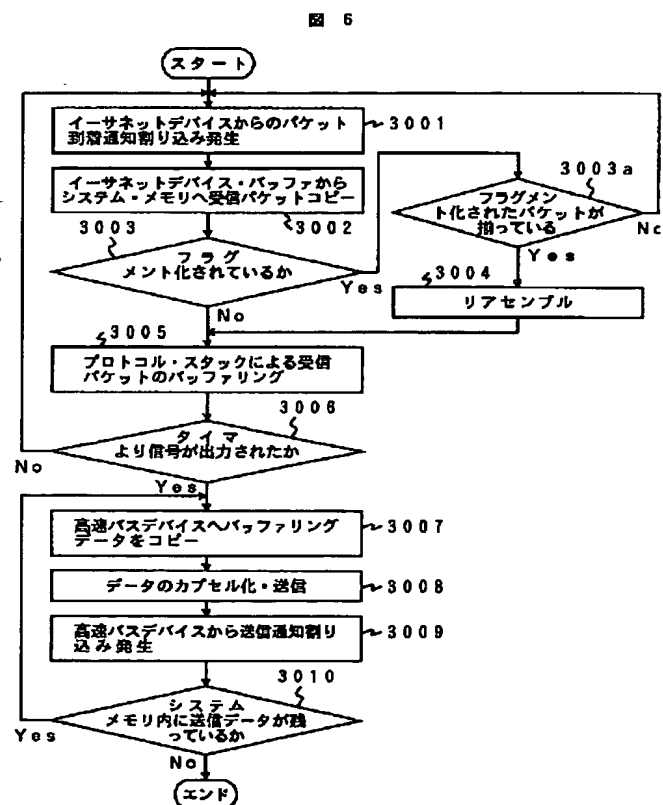
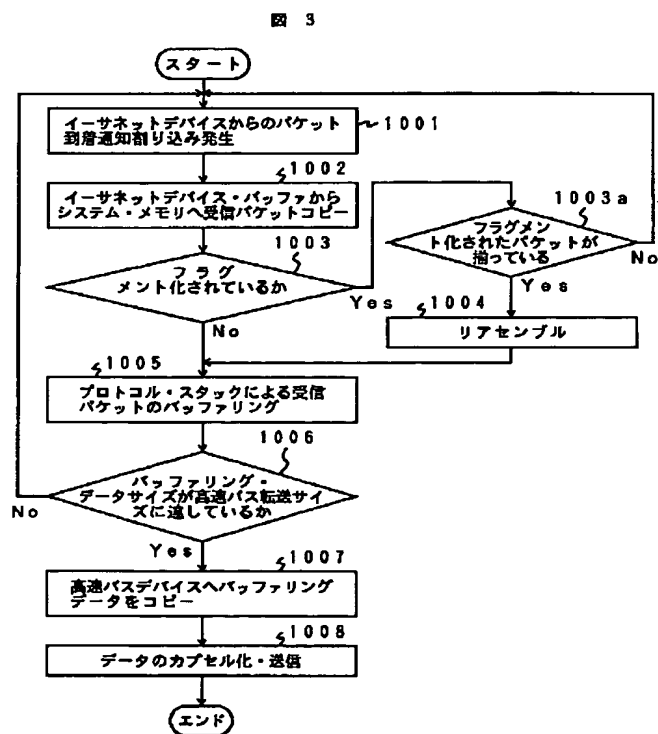


【図5】



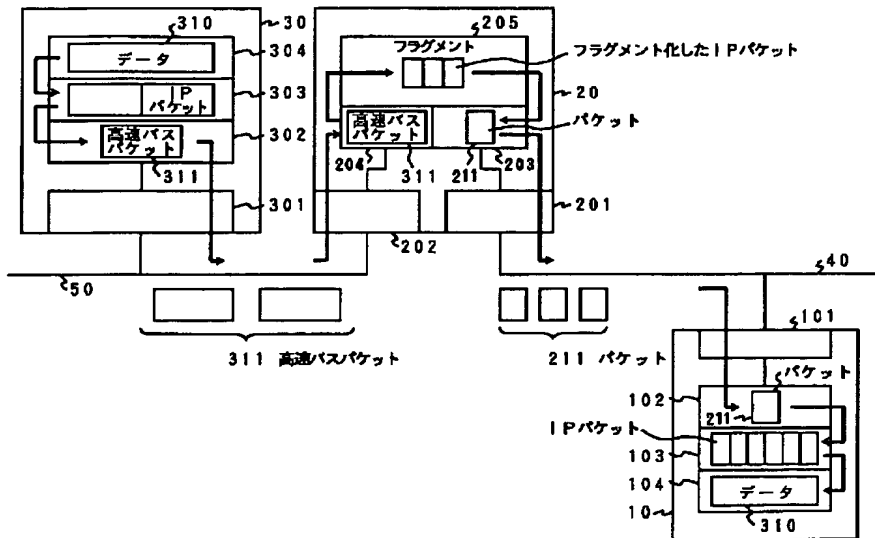


【图 6】



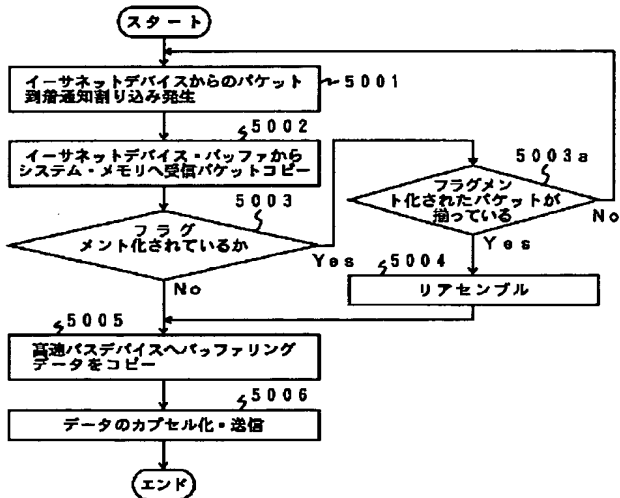
【図4】

図4



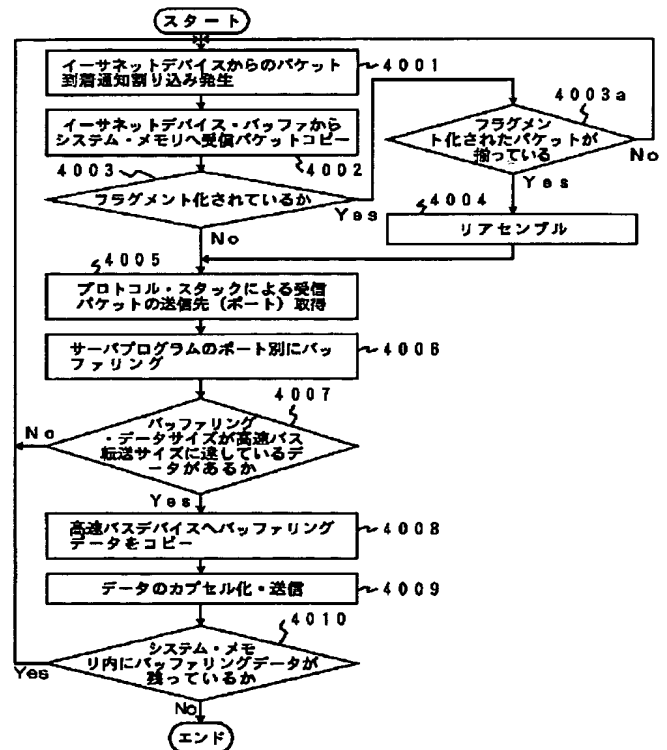
【図7】

図7



【図8】

図8



フロントページの続き

(72)発明者 竹内 理  
神奈川県川崎市麻生区王禅寺1099番地 株  
式会社日立製作所システム開発研究所内

(72)発明者 中野 隆裕  
神奈川県川崎市麻生区王禅寺1099番地 株  
式会社日立製作所システム開発研究所内